

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1-13. (Cancelled)

14. (Currently Amended) A method of associating a phenotype with one or more candidate chromosomal regions in a genome of a species using a phenotypic data structure that comprises a difference in a phenotype between different strains of said species, said genome including a plurality of loci, said method comprising:

establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species;

determining a correlation value for said genotypic data structure ~~based on~~ by a comparison of said phenotypic data structure with said genotypic data structure; and

repeating said establishing and determining steps for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined; wherein

each element in each respective genotypic data structure in said one or more genotypic data structures corresponds to a difference of at least one component of said locus corresponding to the respective genotypic data structure; wherein, for each element in each

respective genotypic data structure in said one or more genotypic data structures, said different strains of said species are selected from a plurality of strains of said species;

an amount that a variation contributes to said at least one component of a locus corresponding to a genotypic data structure in said plurality of genotypic data structures is a function of a distance said variation is away from a center of the locus; and

wherein a genotypic data structure in said plurality of genotypic data structures comprises a plurality of variations that are distributed about the center of a locus corresponding to the genotypic data structure, and said establishing step further comprises:

fitting a distribution of said plurality of variations about the center of said locus with a probability function; and

weighting each variation by a corresponding value derived from said probability function such that variations further from the center of said locus are downweighted so that they contribute less to said genotypic data structure than loci that are closer to said center of said locus.

15. (Original) The method of claim 14 wherein said probability function is a Gaussian probability distribution, a Poisson distribution, or a Lorentzian distribution.

16. (Cancelled)

17. (Currently amended) A method of associating a phenotype with one or more candidate chromosomal regions in a genome of a species using a phenotypic data structure that comprises a difference in a phenotype between different strains of said species, said genome including a plurality of loci, said method comprising:

establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species;

determining a correlation value for said genotypic data structure ~~based on~~ by a comparison of said phenotypic data structure with said genotypic data structure; and

repeating said establishing and determining steps for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined; wherein,

said correlation value is formed in accordance with the expression:

$$c(P, G^L) = \frac{\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

18-19. (Cancelled)

20. (Currently amended) A method of associating a phenotype with one or more candidate chromosomal regions in a genome of a species using a phenotypic data structure that comprises a difference in a phenotype between different strains of said species, said genome including a plurality of loci, said method comprising:

establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species;

determining a correlation value for said genotypic data structure ~~based on~~ by a comparison of said phenotypic data structure ~~to~~ with said genotypic data structure; and

repeating said establishing and determining steps for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective

genotypic data structure in the plurality of genotypic data structures, ~~an associated~~
determining a correlation value;

identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined; wherein,

said correlation value is formed in accordance with the expression

$$c(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)] \times Z}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure;

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure; and

Z is a function of a number of components in the locus, corresponding to the genotypic data structure, having a variation between different strains of said species.

21. (Original) The method of claim 20, wherein said function is selected from the group consisting of taking the square root of Z , squaring Z , raising Z by the power of a positive integer, taking a logarithm of Z , and taking an exponential of Z .

22. (Currently amended) A method of associating a phenotype with one or more candidate chromosomal regions in a genome of a species using a phenotypic data structure that comprises a difference in a phenotype between different strains of said species, said genome including a plurality of loci, said method comprising:

establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species;

determining a correlation value for said genotypic data structure base on a comparison of said phenotypic data structure with said genotypic data structure; and

repeating said establishing and determining steps for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an assoiated~~ determining a correlation value;

identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures has the property that the correlation value associated with the respective genotypic data structure forms is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined; wherein,

said correlation value is a correlative measure cm that is computed in accordance with the expression:

$$cm(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)]}{\{[\sum^i (p(i) - \langle P \rangle)^2]\}^{1/2}}$$

where,

$cm(P, G^L)$ is said correlative measure;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

23-38. (Cancelled)

39. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a genotypic database for storing variations in genomic sequences of a plurality of strains of a species;

a phenotypic data structure, said phenotypic data structure comprising a difference in a phenotype between different strains of said species; and

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said species, said genome including a plurality of loci, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species stored in said genotypic database;

instructions for determining a correlation value for said genotypic data structure ~~based on~~ by a comparison of said phenotypic data structure ~~to~~ with said genotypic data structure; and

instructions for repeating said instructions for establishing and said instructions for determining, for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

instructions for identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structures is~~ a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined; wherein,

a genotypic data structure in said plurality of genotypic data structures comprises a plurality of variations that are distributed about the center of a locus corresponding to the genotypic data structure, and said instructions for establishing further comprise:

instructions for fitting a distribution of said plurality of variations about the center of said locus with a probability function; and

instructions for weighting each variation by a corresponding value derived from said probability function such that variations further from the center of said locus are downweighted so that they contribute less to said genotypic data structure than loci that are closer to said center of said corresponding locus.

40. (Original) The computer program product of claim 39 wherein said probability function is a Gaussian probability distribution, a Poisson distribution, or a Lorentzian distribution.

41. (Cancelled)

42. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a genotypic database for storing variations in genomic sequences of a plurality of strains of a species;

a phenotypic data structure, said phenotypic data structure comprising a difference in a phenotype between different strains of said species; and

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said species, said genome including a plurality of loci, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species stored in said genotypic database;

instructions for determining a correlation value for said genotypic data structure based on ~~an~~ by a comparison of said phenotypic data structure ~~to~~ with said genotypic data structure; and

instructions for repeating said instructions for establishing and said instructions for [[,]] determining, for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

instructions for identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined; wherein, said instructions for determining include instructions for forming said correlation value in accordance with the expression:

$$c(P, G^L) = \frac{\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

43- 44. (Cancelled)

45. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a genotypic database for storing variations in genomic sequences of a plurality of strains of a species;

a phenotypic data structure, said phenotypic data structure comprising a difference in a phenotype between different strains of said species; and

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said species, said genome including a plurality of loci, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species stored in said genotypic database;

instructions for determining a correlation value for said genotypic data structure ~~based on~~ by a comparison of said phenotypic data structure ~~to~~ with said genotypic data structure; and

instructions for repeating said instructions for establishing and said instructions for determining, for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

instructions for identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined, wherein

said instructions for determining include instructions for forming said correlation value in accordance with the expression:

$$c(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)] \times Z}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;
 $\langle P \rangle$ is a mean value of all elements in said phenotypic data structure;
 $\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure; and
 Z is a function of a number of components in the locus corresponding to the genotypic structure, having a variation between different strains of said species.

46. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a genotypic database for storing variations in genomic sequences of a plurality of strains of a species;

a phenotypic data structure, said phenotypic data structure comprising a difference in a phenotype between different strains of said species; and

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said species, said genome including a plurality of loci, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species stored in said genotypic database;

instructions for determining a correlation value for said genotypic data structure based on ~~an~~ by a comparison of said phenotypic data structure with said genotypic data structure; and

instructions for repeating said instructions for establishing and said instructions for determining, for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

instructions for identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that

correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined, wherein said correlation value is weighted by a number of components in the locus corresponding to the genotypic data structure.

47. (Currently amended) A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a genotypic database for storing variations in genomic sequences of a plurality of strains of a species;

a phenotypic data structure, said phenotypic data structure comprising a difference in a phenotype between different strains of said species; and

a program module for associating a phenotype with one or more candidate chromosomal regions in a genome of said species, said genome including a plurality of loci, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species stored in said genotypic database;

instructions for determining a correlation value for said genotypic data structure ~~based on~~ by a comparison of said phenotypic data structure with said genotypic data structure; and

instructions for repeating said instructions for establishing and said instructions for determining, for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an associated~~ determining a correlation value;

instructions for identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that

correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined, wherein,

said instructions for determining include instructions for forming said correlation value in accordance with a correlative measure cm that is computed in accordance with the expression:

$$cm(P, G^L) = \frac{[\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)]}{\{[\sum^i (p(i) - \langle P \rangle)^2]\}^{1/2}}$$

where,

$cm(P, G^L)$ is said correlative measure;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

48-57. (Cancelled)

58. (Currently amended) A computer system for associating a phenotype with one or more candidate chromosomal regions in a genome of a species, said genome including a plurality of loci, the computer system comprising:

a central processing unit;

a memory, coupled to the central processing unit, the memory storing:

a genotypic database for storing variations in genomic sequences of a plurality of strains of said species;

a phenotypic data structure that comprises a difference in a phenotype between different strains of said species; and

a program module, said program module comprising:

instructions for establishing a genotypic data structure, said genotypic data structure corresponding to a locus in said plurality of loci, said genotypic data structure comprising a variation of at least one component of said locus between different strains of said species stored in said genotypic database;

instructions for determining a correlation value for said genotypic data structure based on ~~an~~ by a comparison of said phenotypic data structure with said genotypic data structure; and

instructions for repeating said instructions for establishing and said instructions for determining, for each locus in said plurality of loci, thereby establishing a plurality of genotypic data structures and, for each respective genotypic data structure in the plurality of genotypic data structures, ~~an-associated~~ determining a correlation value;

instructions for identifying one or more genotypic data structures in said plurality of genotypic data structures, wherein the correlation value for each respective genotypic data structure in the one or more genotypic data structures ~~has the property that the correlation value associated with the respective genotypic data structure forms~~ is a high correlation value relative to the correlation values of genotypic data structures in said plurality of genotypic data structures that are not in said one or more genotypic data structures; wherein the loci that correspond to said one or more genotypic data structures represent said one or more candidate chromosomal regions that associate with said phenotype and wherein an amount of said genome that is included in each locus in said plurality of loci is predetermined, wherein

said instructions for determining include instructions for forming said correlation value in accordance with the expression:

$$c(P, G^L) = \frac{\sum^i (p(i) - \langle P \rangle) (g(i) - \langle G^L \rangle)}{\{[\sum^i (p(i) - \langle P \rangle)^2] [\sum^i (g(i) - \langle G^L \rangle)^2]\}^{1/2}}$$

where,

$c(P, G^L)$ is said correlation value;

$p(i)$ is a value of the i^{th} element of said phenotypic data structure;

$g(i)$ is a value of the i^{th} element of said genotypic data structure;

$\langle P \rangle$ is a mean value of all elements in said phenotypic data structure; and

$\langle G^L \rangle$ is a mean value of all elements in said genotypic data structure.

59-77. (Cancelled)